

## 三碱基体的模型和能力学研究\*

刘次全 白春礼<sup>①</sup> 王三山<sup>②</sup> 王莹  
荣茂森<sup>③</sup> 祝晓清<sup>②</sup> 黄京飞

(中国科学院昆明动物研究所细胞和分子进化开放研究实验室 昆明 650223)

(云南大学现代生物学中心 昆明 650091)

**摘要** 三碱基体是形成三链 DNA 结构的基元, 对其进行模型能力学研究具有一定的意义。本文在理论上给出了 22 种三碱基体, 其中有 16 种已见文献报道。通过构型能量优化、进而作总能量和氢键能的计算, 比较了这些三碱基体的相对稳定性和氢键能。

**关键词:** 三碱基体, 构象能计算, Watson-Crick 和 Hoogsteen 配对, 稳定性和氢键分析

如果从 1957 年 Davies 等, 首次提出三链核糖核酸的概念算起, 似乎可以说有关三碱基体的问题已经知道大约 30 多年了 (Moffat, 1991)。近几年来, 随着三链 DNA 结构研究的进展或“终于走向成熟” (Moffat, 1991), 文献报道的由 G、C、A、T 4 种碱基形成的三碱基体已超过 15 种 (Lee 等, 1979; Michel 等, 1989; Griffin 等, 1989; Strobel 等, 1990; Htun 等, 1989; Sklenar 等, 1990; Pilch 等, 1990; Been 等, 1991; Beal 等, 1991; Moffat, 1991; Roman 等, 1991; Schultz 等, 1991; Xodo 等, 1991; Michel 等, 1992; Chastain 等, 1992; Donner 等, 1992)。

在一个三碱基体中, 通常同时存在着两种基本的配对方式, 即 Watson-Crick 氢键碱基配对和 Hoogsteen (1959) 所提出的碱基配对, 即 Hoogsteen 氢键碱基配对。此外, 也存在着一些其它的配对情况, 如在 AAA、AAC、AAC<sup>+</sup> 和 AAT 等三碱基体中就存在着两个 Hoogsteen 氢键碱基配对以及以下可能的排列 (Pullman 等, 1969):

Watson-Crick-reversed Hoogsteen

Reversed Watson-Crick-Hoogsteen

Reversed watson-Crick-reversed Hoogsteen

近几年来, 三链 DNA 的研究受到人们越来越多的注视。《Science》等核心杂志近 5 年来接连发表了数目可观的有关三链 DNA 的研究论文和报告, 这也为我们在理论上

\* 中国科学院“八五”重大科研项目内容。 ① 中国科学院化学研究所。

② 中国科学院上海生物化学研究所。 ③ 云南省计算中心。

本文 1993 年 4 月 30 日收到。

探索三碱基体提供了宝贵的资料。

## 模 型 和 计 算

从 Watson-Crick 氢键碱基配对和 Hoogsteen 氢键碱基配对考虑, 在理论上尝试列出图 1 所示的, 不包括碱基互变异构体参与形成的三碱基体在内的 32 种三碱基体。这些三碱基体中, 有一些迄今尚未被实验所证实, 有的甚至在实际上可能是不存在的。

CAC <sup>+</sup>	CAC	CAA	CAT	CAG	CGA	CGT	A <sup>+</sup> GC
CGC <sup>+</sup>	CGC	C <sup>+</sup> AA	C <sup>+</sup> AT	C <sup>+</sup> AG	CGG	C <sup>+</sup> GT	A <sup>+</sup> GG
TGA	TGT	TAA	TAT	C <sup>+</sup> GA	C <sup>+</sup> GG	TAG	A <sup>+</sup> GT
GAA	GAG	AAA	GGA	AGA	GGG	TGG	A <sup>+</sup> GA

图 1 在理论上组合成的三碱基体

Fig. 1 Combinatory base triplets in theory

在处理时, 对于图 1 中的每一个三碱基体, 均按 Watson-Crick 氢键和 Hoogsteen 氢键将三个碱基定在一个共平面上。参照 Arnott 等 (1974, 1976) 的文献和 Fasman 主编的《Handbook of Biochemistry and Molecular Biology》(1976) 中的有关数据确定初始原子坐标。进而在 SGI 微机工作站上运用共轭梯度法作能量极小化的构型优化。每一个三碱基体均优化至能量曲线下降至与 X 轴靠近并平行于 X 轴时为止。可以看到, 在 22 个三碱基体中, 除 TGC 优化 40 步外, 其余均优化 50 步, 能量计算包括:

Internal	Nonbond
Bond	Van der Waals
Angles	Electrostatic
Torsions	Hbonds
Inversions	
Total energy	

## 结 果 和 讨 论

在图 2 中列出了 22 个三碱基体的结构。选择这些三碱基体的原则, 首先是考虑将已见于文献的三碱基体包括在内。其次, 也包括了 AAA、GGG 和 UAU 等三碱基体。

从图 1 和图 2 不难看出, 在三碱基体中同时存在 Watson-Crick 氢键碱基配对和 Hoogsteen 氢键碱基配对的情况下, 位于中间的碱基均系嘌呤碱基 (如 TAT 和 C<sup>+</sup>GC 等大多数三碱基体), 因为嘌呤碱基可以同时提供两个形成氢键的区域。然而, 在一些文献中 (Griffin 等, 1989; Michel 等, 1989; Beal 等, 1991; Been 等, 1991; Roman 等, 1991; Michel 等, 1992) GTG、GCA、AAT、CGT 和 GGC 三碱基体却存

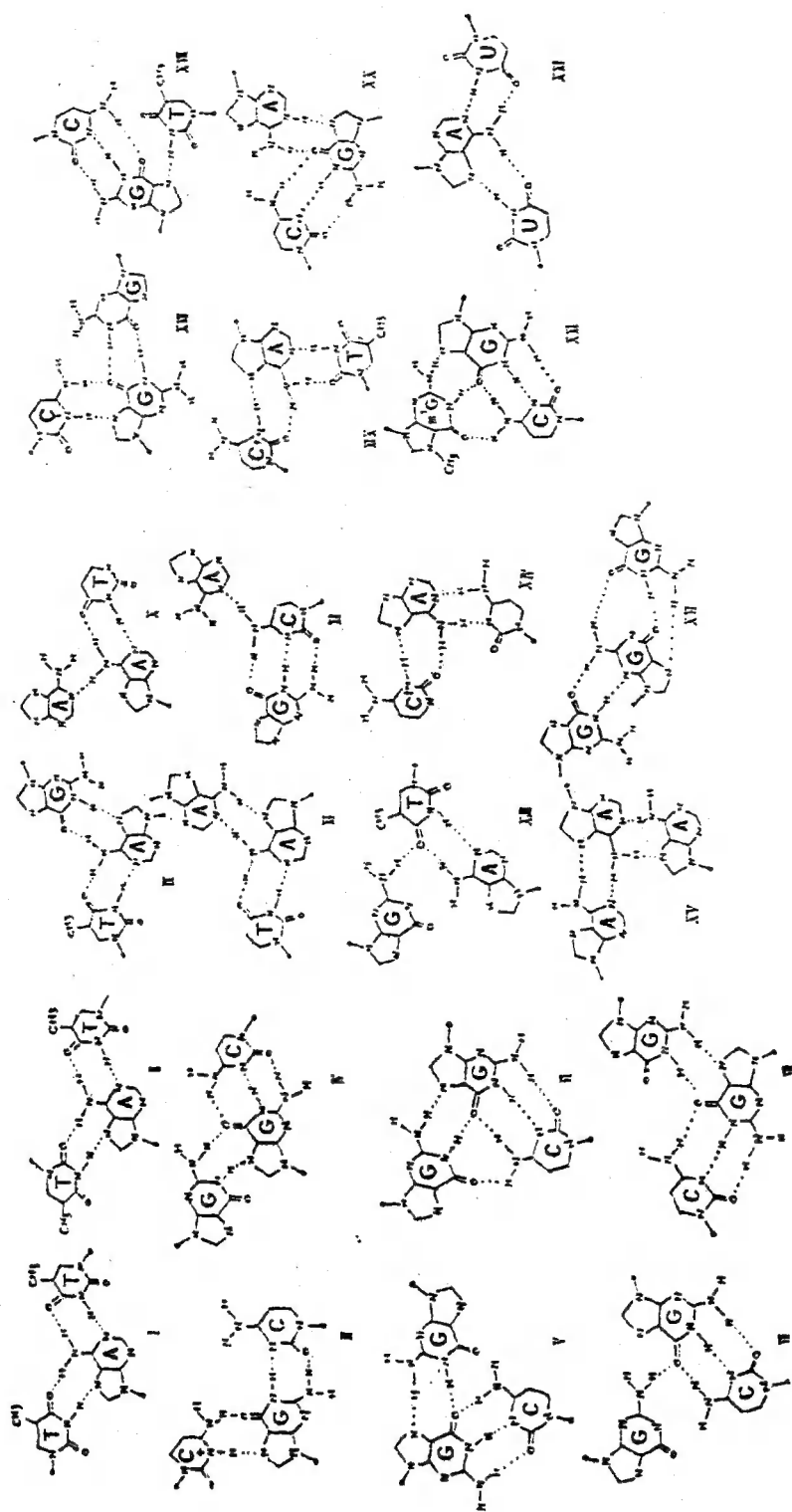


图 2 22 个三碱基模型

Fig. 2 Model of 22 base triplets

探索三碱基体提供了宝贵的资料。

## 模 型 和 计 算

从 Watson-Crick 氢键碱基配对和 Hoogsteen 氢键碱基配对考虑, 在理论上尝试列出图 1 所示的, 不包括碱基互变异构体参与形成的三碱基体在内的 32 种三碱基体。这些三碱基体中, 有一些迄今尚未被实验所证实, 有的甚至在实际上可能是不存在的。

CAC <sup>+</sup>	CAC	CAA	CAT	CAG	CGA	CGT	A <sup>+</sup> GC
CGC <sup>+</sup>	CGC	C <sup>+</sup> AA	C <sup>+</sup> AT	C <sup>+</sup> AG	CGG	C <sup>+</sup> GT	A <sup>+</sup> GG
TGA	TGT	TAA	TAT	C <sup>+</sup> GA	C <sup>+</sup> GG	TAG	A <sup>+</sup> GT
GAA	GAG	AAA	GGA	AGA	GGG	TGG	A <sup>+</sup> GA

图 1 在理论上组合成的三碱基体

Fig. 1 Combinatory base triplets in theory

在处理时, 对于图 1 中的每一个三碱基体, 均按 Watson-Crick 氢键和 Hoogsteen 氢键将三个碱基定在一个共平面上。参照 Arnott 等 (1974, 1976) 的文献和 Fasman 主编的《Handbook of Biochemistry and Molecular Biology》(1976) 中的有关数据确定初始原子坐标。进而在 SGI 微机工作站上运用共轭梯度法作能量极小化的构型优化。每一个三碱基体均优化至能量曲线下降至与 X 轴靠近并平行于 X 轴时为止。可以看到, 在 22 个三碱基体中, 除 TGC 优化 40 步外, 其余均优化 50 步, 能量计算包括:

Internal	Nonbond
Bond	Van der Waals
Angles	Electrostatic
Torsions	Hbonds
Inversions	
Total energy	

## 结 果 和 讨 论

在图 2 中列出了 22 个三碱基体的结构。选择这些三碱基体的原则, 首先是考虑将已见于文献的三碱基体包括在内。其次, 也包括了 AAA、GGG 和 UAU 等三碱基体。

从图 1 和图 2 不难看出, 在三碱基体中同时存在 Watson-Crick 氢键碱基配对和 Hoogsteen 氢键碱基配对的情况下, 位于中间的碱基均系嘌呤碱基 (如 TAT 和 C<sup>+</sup>GC 等大多数三碱基体), 因为嘌呤碱基可以同时提供两个形成氢键的区域。然而, 在一些文献中 (Griffin 等, 1989; Michel 等, 1989; Beal 等, 1991; Been 等, 1991; Roman 等, 1991; Michel 等, 1992) GTG、GCA、AAT、CGT 和 GGC 三碱基体却存

结构形成的基础,那么,三碱基体中碱基间的相互作用,则可看作是形成 DNA 三链结构的基础。

关于碱基间的配对相互作用,王莹等(1990)曾作过理论计算,并对“氢键专一相互作用”的概念提出了异议。最近,又完成了对三碱基体中碱基间相互作用的理论研究(另文发表)。

通过对 22 个三碱基体所作的模型和力学的理论研究,获得了不少有价值的信息和数据。当然,本文的工作仅仅是初步的,还有待于实验研究和更深入的理论研究的检验。



图 3 GGG 三碱基体经构型优化后的结构模型

Fig. 3 Structural model of GGG after Configurational Optimization

## 参 考 文 献

- 王莹,刘次全,黄京飞. 1990. 生物化学与生物物理学报, **22**: 265—270.
- Arnott S *et al.* 1974. *J. Mol. Biol.*, **88**: 509.
- Arnott S *et al.* 1976. *Nucleic Acids Res.*, **3**: 2459.
- Beal P A, Dervan P B. 1991. *Science*, **251**: 1360.
- Been M D, Perronnta A T. 1991. *Science*, **252**: 434.
- Chastain M, Tinocok Jr I. 1992. *Nucleic Acids Res.*, **20**: 315.
- Donner C, Kurlnand M. 1992. *Mol. Gen. Genet.*, **115**: 49.
- Fasman G D. 1976. *Handbook of Biochemistry and Molecular Biology*. Ohio: CRC Press.
- Griffin L C, Dervan P B. 1989. *Science*, **245**: 967.
- Hoogsteen K. 1959. *Acta Cryst*, **12**: 822.
- Htun H, Dahlberg J E. 1989. *Science*, **243**: 1571.
- Lee J S *et al.* 1989. *Nucleic Acids Res.*, **6**: 3073.
- Michel D *et al.* 1989. *Nature*, **342**: 391.
- Michel D *et al.* 1992. *Nucleic Acids Res.*, **20**: 439.
- Moffat A S. 1991. *Science*, **252**: 1374.
- Pilch D S *et al.* 1990. *Proc. Natl. Acad. Sci., USA* **87**: 1942.
- Pullman B, Pullman A. 1969. *Prog Nucleic Acid Res, Mol Biol.*, **9**: 327.
- Roman R F *et al.* 1991. *Science*, **254**: 270.
- Schultz S C *et al.* 1991. *Science*, **253**: 1001.
- Sklenar V, Feigon J. 1990. *Nature*, **345**: 836.
- Strobel S A, Dervan P B. 1990. *Science*, **249**: 73.

Xodo L E *et al.* 1991. *Nuclei Acids Res*, 19: 5625.

## The Investigation of Triplet-base Model and Energetics \*

Liu Ciquan Bai Chunli<sup>①</sup> Wang Sanshan<sup>②</sup> Wang Ying

Rong Maosen<sup>③</sup> Zhu Xiaoqing<sup>②</sup> Huang Jingfei

*(Laboratory of Cellular and Molecular Evolution,*

*Kunming Institute of Zoology, Academia Sinica, Kunming 650223)*

*(Modern Biological Center, University of Yunnan, Kunming 650091, China)*

It is significant to study the model and energetics of the triplet-base, an element of forming three-stranded DNA structure. This paper, at theory, presented 22 kinds of triplet-bases, 16 of which have been reported in references. The relative stability and hydrogen bond energy of these triplet-bases have been compared by optimizing their configurational energy and computing the total and hydrogen bond energy.

**Key words:** Triplet-base, Computing configurational energy, Watson-Crick and Hoogsteen pairs, Stability and hydrogen bond analysis

---

\* This project was supported by the Great Project of the Chinese Academy of Sciences.

① Institute of Chemistry, Academia Sinica.

② Shanghai Institute of Biochemistry, Academia Sinica.

③ The Computing Center of Yunnan Province.